5-1-2010

Fourth and Fifth Grade Student Achievement and Possible Relationship to Three Teacher Selection Tools

Jon D. Hall
University of Louisville and Western Kentucky University, schoolhr@directv.net

FOURTH AND FIFTH GRADE STUDENT ACHIEVEMENT AND POSSIBLE RELATIONSHIP TO THREE TEACHER SELECTION TOOLS

By

Jon D. Hall
B.S., Western Kentucky University, 1974
M.S., Western Kentucky University, 1983

A Dissertation
Submitted to the Faculty
of the Graduate Schools of the University of Louisville
and Western Kentucky University
in Partial Fulfillment of the Requirements
for the Degree of

Doctor of Philosophy

Department of Educational Leadership, Foundations, and Human Resource Education
University of Louisville
Louisville, Kentucky

Department of Educational Administration, Leadership, and Research
Western Kentucky University
Bowling Green, Kentucky

May 2010

FOURTH AND FIFTH GRADE STUDENT ACHIEVEMENT AND POSSIBLE RELATIONSHIP TO THREE TEACHER SELECTION TOOLS

By

Jon D. Hall
B.S., Western Kentucky University, 1974
M.S., Western Kentucky University, 1983

A Dissertation Approved on

April 8, 2010

by the following Dissertation Committee:

_____
Dissertation Director

_____

_____

_____

_____

DEDICATION

This dissertation is dedicated to my amazing wife, Debbie Birdwell Hall and to my wonderful family. Thanks for not giving up on me, constantly supporting me, and allowing me the flexibility to trudge along and get this dissertation completed.

ACKNOWLEGDEMENTS

In working through this research study and resulting dissertation, I have so many people to thank for their cooperation and help. I want to thank my dissertation committee that has helped me make this a finished product and to all the other people in the Cooperative Doctoral program at the University of Louisville and at Western Kentucky University that assisted me with this process.

To my long time mentor in Educational Leadership, Dr. Chris Wagner, no person has been more important to educational success than you. Words alone cannot express all my gratitude for sticking with me over the years to be sure I did not give up before I completed this educational journey. You boosted me over every obstacle (and there were plenty) I met during the research portion of this process

I want to thank all the teachers that cooperated with me in this research, you have bolstered my belief that teachers are the bedrock of public education and you are the most important factor in student learning. Thank you to the superintendents and central office personnel that assisted me in communication, coordination, and collection of data, because I could only successfully gather this data with your help.

I need to give a special thank you to Jeanelle McGuire for being the very best collaborative partner I had during this research project. Your assistance was invaluable and provided the bulk of reliable data that I gathered in this study.

ABSTRACT

FOURTH AND FIFTH GRADE STUDENT ACHIEVEMENT AND POSSIBLE RELATIONSHIP TO THREE TEACHER SELECTION TOOLS

Jon D. Hall

May 8, 2010

This study attempts to identify possible relationships between cognitive ability, teacher dispositions, and content knowledge and elementary classroom achievement. Data was gathered on teachers using the Wonderlic Personnel Test™, Star Teacher Pre-Screener™, and Praxis II™ assessments and compared to classroom growth on the Measures of Academic Progress (MAP) computer-based adaptive assessment. The sample size was limited but included fourth and fifth grade teachers that volunteered from selected Kentucky public school districts. Pearson correlation analysis did not show any significant relationship between the scores on the three teacher assessments and student academic growth using the MAP growth index scores for corresponding classrooms. Implications of this study are that although no relationship was shown, more research is needed that would include larger sample populations and refined study protocol to avoid the limitations noted in this study.

TABLE OF CONTENTS

CHAPTER I

INTRODUCTION

Background

Parents are the first and most influential teachers in a child's life, and most parents realize the tremendous importance of this responsibility. This is easily exhibited in the excitement of a parent teaching a baby to perform functions such as talking and walking. For the first few years of a child's life, the parent is the primary educator. It is not surprising then that turning this responsibility over to another adult can be fraught with trepidation. However, most parents agree that the job of educating their children is a responsibility to be shouldered by government through a system of public schools. Teachers, who ultimately take over the major educational role, will undoubtedly need to prove their worth to the parents of our children. No responsible parent would intentionally choose to place their child with a teacher who they know will do a poor job teaching.

A cursory retrospective of education, even in ancient times, describes the basic needs of society and supports the establishment of an educational system. It is sufficient to declare that the basic needs of society led to the establishment of public education. Religious institutions in Europe led the development of education for centuries as a means to create, preserve, and transmit human culture. Since most of the population of the new world had little need for a formal education, early education was designed for and carried out by white, wealthy male land-owners. By the time of the American

Revolution, this responsibility was transferred to communities and eventually to government agencies where “The writings of European theorists were eagerly read, new institutions sprang up, noted American leaders formulated distinctive theories of education and the beginnings of a system of public education emerged” (Eby & Arrowood, 1934, p. 530).

Later, the establishment of a free public education for all led to the dilemma of selecting the best person to teach disparate of children from a wealth of diverse cultures and family units. In the past, selection of teachers usually meant attention to one or more of the following criteria:

> (a) an assessment of the prospective teacher's character, values, and beliefs, usually by a member of the clergy; (b) an assessment of the prospective teacher's knowledge in selected domains, usually by a common or standardized test in the teaching subject; and/or (c) an assessment by a faculty with regard to the prospective teacher's course of professional study, usually with a major emphasis on pedagogy and teaching skill (Murray, 2007).

Most people understood that the development of high quality craftsmen in the trades involved the transfer of knowledge from the master to the apprentice. This assumption was adopted without challenge because an apprentice desiring to become highly skilled would want to learn from the most accomplished craftsman. In like manner, the first school teachers may have been the most educated members of the community willing to accept the chore of teaching the skills of reading, writing, and arithmetic to the children of the community. Logically, it follows that the person selected to teach in these new institutions should be the most intelligent person available.

Defining education as a simple transfer of knowledge from teacher to student might have been appropriate prior to the last half of the twentieth century when knowledge was limited and could be well covered in textbooks. We may find that direct transfer of know-how may still be effective when training in a physical skill, but late in the twentieth century, information became so vast and general knowledge was increasing at such a high rate that no one person could be completely knowledgeable in a subject area. Many people agree that intellectual knowledge alone is not a guarantee of success in teaching and can give personal accounts of great teachers that did so much more than just transfer knowledge to them.

The education of our young is a critical part of contemporary society. Front and center in the paradigm of public education is the task of recruiting and retaining high quality teachers. The need for effective teachers is more critical today than ever before as the need for dramatic change in American education was introduced as a public topic, thanks in part to the 1983 publication of "A Nation at Risk" by the National Commission on Excellence in Education. The report bemoaned the fact that American students are behind many other developing countries and that we must improve the educational opportunities of our young people in order to compete in the new global economy. On March 31, 1994, President Clinton signed Public Law 103-22, commonly referred to as the "Goals 2000: Educate America Act," which set forth a blueprint for educational reform to reverse this trend (Knudsen & Morrissette, 1998).

Societal changes during the twentieth century caused a transformation in the American family from the norm of a two-parent family with a stay-at-home mother to more variable family compositions. The movement of women from the home into the

workforce created the need for child-care very early in children's life with a large majority of American mothers utilizing child-care in order to stay in the workforce. According to Deborah Phillips, Study Director for the Board on Children, Youth and Families of the National Research Council, "the typical infant in the United States is in child care prior to the age of five months and is in care more than 30 hours a week" (Child Care and Early Learning, 1999). The need for better child care programs with trained personnel and appropriate facilities is being met by public agencies across the United States. From California to Connecticut, and most states in between, early childhood initiatives can be found in diverse programs utilizing funding from sources often tied to lotteries, tobacco, or alcohol. This indicates that the public recognizes the critical need for effective teachers from early childhood through graduation from high school (Child Care and Early Learning, 1999).

The Teacher Quality Problem

Given the high level of public concern for the need to improve student achievement in America, it is understandable that the public has voiced concern about the quality of teachers in our public schools. Teacher quality may be defined in many ways depending on the outlook of the definer. However, general agreement exists that teachers of high quality are those who get their students to learn and develop in positive ways at high levels. As pointed out by Goe (2007), disagreement abounds in what teacher characteristics are most important in getting high levels of achievement from students. Educational researchers have approached this issue from many paradigms, emphasizing different aspects of teacher quality and effectiveness according to their own research goals (p 4).

The book, *Qualities of Effective Teachers*, shows that the study of effective teachers has revealed the presence of abilities and talents that enable them to deeply engage their students in hands-on learning activities. Research shows that teachers with a caring personality can motivate students to learn even the most challenging material. A caring personality is more than just knowing the students and involves personality traits such as patience, trust, honesty, and courage. Attributes of caring include being a good listener, exhibiting gentleness and understanding, as well as developing individual knowledge of students on a personal level. Effective teachers have been shown to have an overall love for children and support them with warmth and encouragement. The best teachers today must be able to teach students how to think and research for themselves in order to be successful (Stronge, 2002).

It is critical to note that no universal agreement exists on the issues of qualities impacting teacher effectiveness. The undergirding premise of these debates is that teachers make a difference in student learning leading to high levels of academic achievement. Researchers disagree on which teacher characteristics are reliable predictors of student success. The debate expands to encompass how and to what level student achievement data should be utilized for the evaluation of teachers, yet limited large scale research has been conducted to calculate the influence teachers have on student achievement using the standardized assessments widely accepted today (Wright, Horn, & Sanders, 1997).

Sanders and Rivers (1996) have shown the enormous difference teachers can make in the achievement of their students. Children assigned to ineffective teachers probably will never regain the lost achievement, even when assigned to very effective

teachers in later years. Furthermore, children assigned to effective teachers for three consecutive years scored 49 percentile points higher on a standardized reading assessment than students assigned to three ineffective teachers (p. 4). Sanders developed the methodology used for this study, the Tennessee Value Added System (TVAAS). TVAAS utilizes statistical mixed model theory and methodology in a complex analysis of longitudinal student achievement and has been used to establish a measure of effectiveness of schools and even individual teachers (Sanders & Horn,1998).

Within this decade, both professional educators and policymakers tend to agree that utilizing student achievement as a measure of teacher quality is an appropriate action due to the high emphasis placed on standardized achievement test scores. The availability of this data has improved due to reliance on widespread state assessments through high-stakes testing for NCLB and state accountability. This naturally leads to use of this data in research on teacher quality (Blanton, Sindelar, & Correa, 2006) and the development of this research project.

Significance of the Study

Today, educational administrators still struggle with the challenge of hiring teachers possessing the skills and knowledge to develop successful students. This is arguably the greatest strategic challenge schools will face during the first couple of decades of the twenty-first century. The competition to hire enough top quality teachers will increase for local school districts. School officials responsible for hiring must consider whether they are doing what is necessary to attract the best teacher candidates compared to competing school districts (Lee, 2005). This problem can be critical in two extreme situations, specifically urban inner-city locations (Jacob, 2007), as well as very

rural schools (Little, 1998).

The focus of this study is how to predict the strength and career success of teachers and is based on the need for objective measures to use in selection of teacher candidates during the hiring process that will result in a high success rate. This study attempts to identify possible relationships between three traditional measures of teacher candidates and student achievement. The three evaluative measures typically used for selection and hiring of new teachers by school districts are cognitive ability, content knowledge, and teaching strengths. Cognitive ability will be measured by the Wonderlic Personnel Test (WPT); content knowledge by the Praxis II™, and teaching strengths by the Star Teacher Pre-Screener (STP)™ assessment. Student achievement will be measured using Northwest Educational Association's assessment, Measures of Academic Progress (MAP) at the fourth-grade and fifth-grade level. This test is given at least twice per year to determine individual student growth over the year.

The issue of hiring the best teachers is not just a philosophical problem, but can carry profound economical impact for school districts. According to the National Commission on Teaching and America's Future publication, *Policy Brief: The High Cost of Teacher Turnover*, the costs related to teacher turnover for 2005 were estimated to be $7.3 billion nationally. Emphasis was on the premise that a primary focus for school leaders should be that of hiring the right teachers and then giving them proper support to increase retention (Barnes, Crowe, & Schaefer 2007).

Costs are associated with both pre-hire and post-hire activities. Consider the expense involved with recruiting, application processing, screening tests, interviewing, and decision-making activities by school personnel in pre-hire activities at the school and

district level. Once a candidate has been selected, the costs continue to mount due to legal and practical requirements prior to placing the teacher in the classroom (Peterson, 2002). Human Resources personnel must ensure that all employment requirements have been met, all credentials and criminal background checks have been confirmed, and the personnel file is complete (Scherer, 2003).

In a five-district study of these costs, the amount per teacher-leaver ranged from $17,872 in Chicago, Illinois, where 25,300 teachers left the district, to $4,366 in Jemez Valley, New Mexico, where only 41 teachers left. These costs were determined by surveys of time and funds expended on activities associated with teachers leaving the districts and post-hire expenses for professional development and separation activities. The annual costs for large districts can be astounding, as determined by the NCTAF Teacher Turnover Cost Calculator found at www.nctaf.orgh, which was used to estimate total costs for some of the larger U.S. districts. These estimates of annual costs ranged from $94,211,250 in Los Angeles, California, to $4,462,500 in Hartford, Connecticut (Barnes et al. 2007).

All expenditures, coupled with the additional training many school districts now provide to implement educational strategies to address student achievement targets, can result in investments of thousands of dollars per new teacher hired. With this type of investment in new teachers during the first year, school districts need a high retention rate to prevent repeating this monetary outlay. In an average school district, if one third of new first-time teachers leave before completing five years (Skandera & Sousa, 2000), hiring an ineffective or mismatched replacement teacher can be a very expensive mistake. It goes without saying that most school districts cannot afford to expend thousands of

dollars for every hire that may or may not be effective in the classroom.

One way school districts could improve teacher retention is to hire the right people. Chances for success can be enhanced by improving screening and selection practices prior to employment using assessments and parameters predictive of teaching success. If school districts could do a better job of predicting teaching success among the pool of applicants deemed to be highly qualified by NCLB before hiring, the negative outcomes, both monetary and academic, could be reduced.

The supposition can be made that school districts are directly comparable to companies and their search for teachers comparable to the search for managerial talent in business. It is important for schools as well as companies to realize that recruitment of qualified individuals is imperative. The looming shortage of talented individuals is comparable to a fight for available talent, and most companies and schools are not well equipped for this competition. In order to thrive, schools, like American companies, must realize that dramatic changes and immediate action are necessary to obtain the talented individuals needed to improve the American educational system (Lee, 2005).

Unfortunately, too many school districts suffer from poorly designed teacher selection systems that have devoted little time and training for the persons responsible for hiring decisions. Researchers have found that limited applicant pools, inadequate screening processes, and lack of organization often plague school districts and result in poor hiring decisions (Peterson, 2002). Hiring ineffective teachers is exactly what most districts do when the pool of candidates gets shallow or there are extreme time constraints to get teachers in the classrooms. As pointed out by Murname and Steele (2007), “School districts often respond to shortage of effective teachers at the prevailing wage not by

leaving teaching positions vacant, but by filling them with ineffective teachers" (p.18). This poor management practice is accentuated by NCLB requirements for highly qualified teachers in every classroom based upon credentials only, with no guarantee of the effectiveness of the teacher.

Acknowledging that great teachers have something more than their intellectual knowledge and cognitive ability working for them in the classroom is the next step in deciding what to look for in potential teacher candidates. As noted by Liesveld and Miller (2005), great teachers are unorthodox in the way they operate, depending on their own instincts and developing strategies that are successful in aiding students to attain their own educational goals. Much of the research on effective teachers during the past four decades indicates that the characteristic that sets great teachers apart from mediocre teachers is individual talent (pp.15-16).

Individual Student Achievement

Public educational systems are held accountable by state and federal government for student achievement by examining performance on high-stakes student tests. These mandated assessments can be a highly controversial issue. As they continue to be used for this purpose, more information should be available concerning the impact on all stakeholders (Vogler, 2002). It is important to note that these high-stakes tests designated as the indicators of the success of schools in teaching core content to students are not valid measures of individual student learning. Smith and Fey (2000) lay out a very convincing argument for the incongruence of using the same assessment process designed for accountability purposes to determine valid student achievement data. Their analysis indicates opposing forces are at work because assessment for accountability arises from

political situations and invariably contain flaws in construct validity due to compromises in testing standards. Furthermore, practices in teaching and testing tend to change in order to promote better scores on the tests without real improvement in student learning, resulting in score inflation. Instead of assisting the students most disadvantaged, the flawed system may put them at greater risk of falling behind (p. 334).

Data accumulated from student testing in systems designed for accountability may not be useful for other purposes, but the inverse situation may be feasible. Using valid student assessment practices that are designed to monitor individual student achievement on a formative basis should allow evaluation of progress toward state or national performance standards.

The desire for longitudinal data to indicate growth in individual student learning led to development of a new genre of testing. These tests are adaptive to the individual student's level of learning in order to obtain a valid profile of ability levels in basic skill areas, and when used at prescribed intervals, will indicate student progress made over time. By aggregating all students in a classroom or at some other level such as an entire grade, school level progress can be determined.

The Northwest Evaluation Association developed a computerized testing process specifically designed for determination of student growth over time. The Measures of Academic Progress (MAP) assessment is computerized, which allows the test to be adaptive each student by altering the difficulty of the items based upon the individual student's performance. The goal of this adaptation is to arrive at an assessment level where the student answers only half of the test items correctly. MAP uses a growth index score which compares a student's growth to national norms of students that started at the

same level expressed as a Rasch Unit (RIT) score (Cronin & Bowe, 2006). Since the MAP is an adaptive test, a RIT growth goal is calculated for each student according to their individual achievement level. A growth index (GI) score is calculated to the individual students after the end-of-year test which indicates the actual variance between the RIT goal and end-of- year score. GI can be used in comparison studies such as this research project that compare the average growth index score for a teacher's students with that teacher's performance level on three common teacher assessments.

Research Questions

The review of literature revealed the current emphasis on examining student achievement and the need to improve teacher quality within our public schools. This research project was designed around three primary research questions, they are:

Research Question 1: Is there a relationship between the academic progress of students in reading and math and the cognitive ability of their teacher as measured by the Wonderlic Personnel Test?

Research Question 2: Is there a relationship between the academic progress of students in reading and math and the rating of teacher attributes as measured by the Star Teacher Pre-Screener™ on-line assessment?

Research Question 3: Is there a relationship between the academic progress of students in reading and math and the teacher's knowledge of pedagogy/subject matter as measured by the Praxis II™ test for elementary teachers?

Definitions

Measures of Academic Progress (MAP) – A computer based adaptive achievement test developed and marketed by the Northwest Evaluation Association. This is the student assessment used by participating districts to gain longitudinal student achievement data for their students.

Growth Index (GI) – Numerical representation for academic achievement on the MAP during the school year. It is the variance of the actual RIT score from the predicted RIT score for students taking the adaptive test Measures of Academic Progress from the Northwest Evaluation Association.

Wonderlic Personnel Test (WPT) – Timed, fifty item, pencil and paper test produced by Wonderlic Inc. designed to quickly measure general intelligence level and used in study to give a quantitative score representing teacher cognitive level.

Star Teacher Pre-Screener (STP) – On-line assessment of ten areas of teacher attitudes and dispositions about working with students, designed by the Haberman Educational Institute to rate teacher candidates in their ability to work in urban classrooms with at-risk students living in poverty. Used in this study to give a quantitative score representing teacher dispositions.

Veteran teachers – Teachers with over ten years teaching experience during the student achievement reporting period.

Mid-veteran teachers – Teachers with between four and nine years teaching experience during the student achievement reporting period.

Non-veteran teachers – Teachers with less than four years teaching experience during the student achievement reporting period.

CHAPTER II

REVIEW OF LITERATURE

The Need for Highly Effective Teachers

The landmark legislation of 2001, commonly referred to as No Child Left Behind (NCLB), fueled an already existing clamor for schools to do a better job educating all of our public school children. The requirements of NCLB that sub-groups of students have the same achievement levels of the majority group, has stimulated much discussion about student achievement gaps in American schools over the last decade. NCLB further demanded that all schools have highly qualified teachers in an effort to improve the quality of teaching in public schools.

In order to be declared "highly qualified," NCLB required that teachers have at least a bachelor's degree, hold full state licensure or certification, and be able to exhibit mastery of the subjects they teach. After several years of work by the states to determine their own definition of "highly qualified" that would satisfy the NCLB requirements for highly qualified teachers, it has basically come down to defining "highly qualified" as being certified in a teacher's assigned field. This does a disservice to education because the requirements seem to have little to do with teacher effectiveness in promoting student learning. We now need more discussion about the gap in teaching quality in American schools within the group of certified teachers (NCTAF, 2008).

Boyd, Goldhaber, Lankford, and Wyckoff (2007) studied the impact of post-secondary education and certification on teacher quality and found an extreme range of qualifications of teachers from school to school. Recent research has identified large disparities when comparing teachers' influence on student achievement and found that students most at risk for failure more often have the least effective teachers due to lack of experience, improper certification, or even poor academic ability. This often results in lower academic gains for some teachers assigned to high poverty schools than their peers in schools with a preponderance of more affluent students (Ascher & Fruchter, 2001).

This problem has the potential to become more severe as massive numbers of teachers become eligible to retire in the next few years. Combine those numbers with the percentage of new teachers who leave teaching within the first three years, and the struggle to improve the quality of America's teaching force becomes further complicated (NCTAF, 2008). In one study typical of the problem, 18 percent of teachers left the teaching profession in their first two years while another six percent transferred to another school district.

In addition to many other benefits, hiring and retaining high quality teachers would benefit schools monetarily by overcoming a class size issue. Reducing class size in fourth and fifth grades may benefit student achievement, but data from the UTD Texas Schools Project indicates that an expensive move to reduce class size by ten students per classroom would give a smaller benefit than supplanting a low quality teacher with a teacher one standard deviation higher. The results of this study show wide variance in the impact on student achievement attributable to the teacher and that much of the disadvantages of low socioeconomic conditions can be overcome by high quality

teaching throughout the early grades. It should be noted that the distinct characteristics of teachers who may be responsible for these gains are not readily distinguishable to researchers. However, sufficient evidence is present to indicate no support for the premise that obtaining a master's degree raises a teacher's effectiveness. This is consistent with prior findings, but it should be noted that Kentucky requires a master's degree or equivalent for continued certification. Another finding of this study indicates that, although experience may positively affect quality of instruction for the first two to three years, very little improvement in teaching can be attributed to experience after those initial years of teaching (Rivkin, Hanushek, & Kain, 2005).

Student learning in schools is profoundly affected by the fact that they are taught in groups rather than as individuals. This profoundly affects the ability of teachers to be effective in promoting student learning because it demands knowledge of best practices that promote learning within the social institution called school. In order to become an effective teacher that can promote the growth of a whole classroom full of students, they must hold dispositions with values and beliefs that support implementation of these best practices (Haberman, 2007).

Other reasons abound that support the premise of hiring the highest quality teachers available, but very little evidence exists that school districts are doing a good job of discerning which teachers will be effective. It remains to be seen if some of the efforts to attain accountability by individual states will change the efforts of school districts in hiring and retaining teachers most likely to promote high levels of student learning. In sum, it is impossible to judge whether school districts' hiring decisions are helped or hindered by the constraint of hiring only certified teachers.

A study of teacher certification requirements by the states verified little uniformity across the states as far as changes in requirements for certification. The researchers found that very little research has been conducted to determine the effectiveness of stringent teacher testing requirements for certification as well as few studies relating coursework in pedagogy with student achievement. Some research has found a weak correlation between pedagogical training and student achievement but none of these studies have established a causal link. Furthermore, considering the large investment that is made in teacher preparation and certification nation-wide, there is a need for evidence that current requirements result in improvement in teacher quality. It is possible that the heavy requirements of teacher preparation and certification might actually depress student success by limiting the supply of effective teachers (Boyd, et al., 2007).

Studies have shown a significant correlation between the quality of teachers and the achievement of their students with an effect that is stronger than socioeconomic status (SES) or other student characteristics. Researchers have examined several possible exemplars of teacher quality including forms of cognitive ability evidenced in college tests such as college entrance exams and pre-service exams. A positive relationship also has been established for verbal ability, subject-matter knowledge, and student achievement. The researchers work also posits that teacher dispositions play an equally positive role in student achievement as that shown by pedagogical and content knowledge. The attempt to quantify teacher dispositions is very difficult because they are an innate part of our personality (Singh and Stoloff, 2007).

Perhaps the most notable research has been conducted in Tennessee by William Sanders and his associates using the Tennessee Value-Added Assessment System (TVAAS), which is the system now being used in there for NCLB calculations through a pilot project for the United States Department of Education. When elementary children had excellent teachers for three consecutive years, their math scores averaged at the 96th percentile, compared to an average score at the 44th percentile on the same test for comparable students having ineffective teachers for three consecutive years. In addition, these student gains in academic achievement resulting from effective teachers do not seem to diminish over time. Also, the effect that low performing teachers have on students tends to persist, with students' academic achievement resisting positive change in subsequent years compared to students not subjected to poor teachers. Conversly, when students had high performing teachers, later academic achievement was enhanced from the experience and could be compounded for each year that the students were taught by a highly effective teacher.

Results very similar to those found in Tennessee were obtained from a study in Texas on younger students' performance in both math and reading. First-grade students having three consecutive high-performing teachers resulted in average gains on the math portion of the Iowa Tests of Basic Skills from the 63rd to the 87th percentile, while peers with comparable abilities but less effective teachers lost ground by dropping 42 percentile points in both math and reading. Furthermore, the Dallas study substantiated the results of Sanders' work by indicating the effects were long-lasting and cumulative (Tucker & Stronge, 2005).

In contrast to conventional beliefs to the contrary, the data show that variation in students' ability levels does not alter the effectiveness of the teacher to a great extent. High-scoring students made lower gains than both the average and the lower-scoring students, allowing for the study of all types of classrooms with comparable results. The conclusions drawn from this work led to recommendations to use student academic data indicating longitudinal measures of academic progress as a valid part of teacher evaluation processes. "Though the debate about whether student achievement data should be used as part of an assessment, evaluation, and accountability system for teachers will assuredly continue the results of this study suggest that *teachers do make a difference* in student achievement" (Wright, Horn, & Sanders, 1997, p. 66).

Partly because of the scarcity of other definitive evidence concerning what constitutes high-quality teaching, policy makers have quickly embraced value-added assessment as a tool to determine which teachers are highly effective. The move from once-per-year student testing for establishment of whether schools have achieved the expectations set by NCLB for adequate yearly progress (AYP) to longitudinal tracking of individual student progress is occurring. The next step now being examined by policy makers is rating teachers' performance based upon how their students perform on these new tests. School administrators should be able to confidently evaluate teachers as highly effective by looking at academic growth of their students during the school year (Imig & Imig, 2006).

Student Achievement Assessment

Over time, educational systems have used many methods to evaluate student learning, with the changes primarily reflecting societal advances. For example, after the

industrial era and adoption of mass production practices, testing moved from oral recitation to written responses due to greater supply of paper and writing utensils. The mass production of lead pencils allowed the invention of standardized achievement tests with the help of psychologists and progressive educators. Today they exist in some form of mandatory test in every state. It is conceivable that the next step in student testing will be computer based.

Standardized achievement tests can be classified based upon the interpretation of student performance. The basic types are norm-referenced, criterion-referenced, and value-added assessments. Norm-referenced assessments are typically administered in a standardized format to students across the country in the same grade level to provide a comparison group for statistical calculations. Norm-referenced testing allows determination of how one student's achievement stands in comparison to the others in the group. Norm-referencing also allows a school district to rate how it compares to other school districts or with another school.

Criterion-referenced assessments deal with student performance compared to standards of mastery for the learning domain being tested. These tests establish what level of knowledge the student has obtained compared to standard expectations and can determine the students' strengths and deficiencies in that domain. These tests are used to evaluate school programs and curriculum more often than individual student achievement (Tucker et al., 2005).

Relatively new to the field of standardized testing, value-added and adaptive testing combine the norm-referenced and criterion models, giving educators a better measure of student learning attributable to certain influences such as school or teacher

effectiveness. Value-added measurements are calculated by comparing student achievement to a norm group and then to themselves by retesting at a later date. Although there are a few different value-added assessment programs in place today, one of the best known is the TVAAS discussed earlier in this chapter (Sanders & Horn, 1998).

In a Phi Delta Kappan/Gallup survey conducted in 2007, 82% of the general public preferred the use of individual student improvement over the school year as an accountability measure instead of the percent of students that pass an end-of-year test. This is reflective of what is currently happening in educational circles as value-added methods are gaining favor among educators and policy makers. The move to this type of assessment is slowly gaining a foothold nationally, with five states gaining permission to use value-added systems for AYP accountability as pilot growth model projects. However, not all educators feel value-added assessments are perfect assessors of student achievement or indicators of gains and call for more work to be done to establish testing validity before widespread acceptance is warranted (Amrein-Beardsley, 2008).

Computers and on-line services have allowed the development of new adaptive testing programs which allow students to test at their present achievement level based upon performance as the test is being administered instead of at the same level as all students in a particular grade or group. The computation of student results establishes an achievement level on a criterion-based learning continuum that is subsequently compared to the norm group of students at their own level. Growth index scores can be calculated using the norm group and differences in performance by the individual student over time (Cronin et al., 2006).

Yeh (2006) calls for more use of the computer-adaptive tests to replace the current end-of-year accountability tests mandated by NCLB. The ability to use computer-adaptive test diagnostic data to specifically meet student needs is not exclusive from the possibility of using them for accountability, showing how a school is promoting growth in student achievement regardless of the starting point. He urgently argues the case for using these tests in place of other accountability testing systems.

> The potential of the standards and testing movement to exert a positive influence on teaching and learning is likely to be undermined to the extent that the following issues are not addressed: lack of diagnostic information, slow reporting of results and inappropriate format, lack of individual growth information, and lack of computer-adaptive features that ensure that students receive test forms at appropriate levels (pp. 519-520).

The increased use of the computer-adaptive tests is plausible and deems serious consideration and supports the choice of the computer-adaptive MAP test for this research study.

NWEA pioneered the development of MAP, a computer-adaptive test being used by many school districts nationally and a rapidly growing number of schools in Kentucky. This review of literature has produced evidence to support the reliability of student growth index scores in determining value-added assessments for students of all ability levels useful for multiple purposes ranging from individual diagnostic adaptation of teaching, growth data for groups of students within schools and districts, or even as part of individual teacher evaluations.

Measuring Cognitive Ability of Teachers

Previously it was suggested that selection of highly intelligent individuals to become teachers was a logical process. However, one might question how much influence the cognitive ability of the teacher has on student learning in the classroom. If the premise that cognitive ability tests can measure a person's ability to analyze situations and choose the appropriate actions to take (Arnold, 2006), then it is important information to know before hiring a teacher, as teachers constantly face changing conditions and situations.

The importance of determination of cognitive ability of teachers is given some attention during the pre-service stage of teacher preparation, since these individuals must meet some minimum level for admission to college and then again through pre-certification testing. However, in Kentucky, as with some other states, all teacher candidates are eligible for hire on a somewhat equal status once they receive initial certification.

One aspect of cognitive ability, verbal ability, has been extensively researched over the last four decades. Results are somewhat inconclusive in determining the effect that verbal ability of the teacher has upon students. The two most often cited studies showing a positive link are the Coleman report in 1966 and research by Ferguson in 1991 using data from studies in Texas and Alabama. Conclusions drawn from these studies and espoused by former U.S. Secretary of Education Paige promote the idea that a very close link exists between cognitive ability and verbal ability of the teacher and learning by their students especially in the early grades.

There are limitations to the Coleman (1966) report to be noted:

> First, the Coleman report makes claims for teachers' verbal ability and student achievement only for African American students, not for the combined student scores. Second, the relationship emerges only in Grades 3 and 6. Third, the limited focus of the instrument the researchers used (vocabulary) casts some doubt on the soundness of making claims about the more complex concept of verbal ability as it applies to teachers (p. 346).

This study was large and encompassed students in 3rd, 6th, 9th, and 12th grades, with the conclusion that verbal ability of the teachers based on a vocabulary test was directly related to student achievement test scores within the sample group (Andrew, Cobb, & Giampietro, 2005).

Testing of teachers developed in a similar manner to that of assessment of students in our public schools. After World War II, the National Teacher Exam emerged but was not widely adopted by states as a factor in deciding which pre-service teachers were ready to teach until the 1970's. The hope for an instrument that would identify effective teachers diminished somewhat, and other measures were adopted to mollify the limitations of these tests of basic skills, pedagogy, and subject matter knowledge (Imig et al., 2006).

The use of teacher exams to quantify readiness of teacher candidates provided a source of data for comparison to student achievement. One such study analyzed the verbal ability of teachers as measured by the Texas Examination of Current Administrators and Teachers (TECAT), which was administered to active teachers and administrators in Texas prior to recertification and compared to math scores of students from the Texas Educational Assessment of Minimum Skills (TEAMS). Odd year grades

from first through eleventh were tested and compared to prior scores from 1985 through 1990. This study indicated that a range between 19-25% of average TEAMS reading score variance was attributed to the teachers' TECAT score with the exception of the first grade. The relationship in this study was somewhat weaker than what is often reported because 75% of the variance was attributed to other factors (Andrew et al., 2005).

Undoubtedly, the school districts have the opportunity to do more evaluation of overall intelligence in their teacher candidates and many districts probably have some measure of cognitive ability through the examination of college transcripts. There are some districts that may do some type of assessment to give more data for that purpose. The use of the Wonderlic Personnel Test (WPT) is a fast, economical way to establish a standard method of comparison of all applicants for a teaching position.

The WPT has been in existence for many years, having been developed by the industrial psychologist, Eldon F. Wonderlic, working with Carl Hovland in 1937. They selected and organized test questions in a manner that produced a valid, reliable test that could be administered easily. The use of this test grew steadily and was used by the United States Navy during World War I for pilot and navigator selection. Use of this test has steadily grown over the decades, receiving acceptance by many through several validation studies over this time period. At one time in the mid 1960's it was used by 55% of U. S. companies with 1600 or more employees (Jones, 1973 p. 600). WPT has been used by many organizations, but probably is best known for use by the National Football League to measure cognitive ability in prospective players (Wonderlic Inc. website, 2008, www.wonderlic.com/about-us.aspx#).

Over the years, many researchers have shown that the WPT is a reliable instrument comparable to longer more complicated assessment instruments such as the Woodcock-Johnson-Revised Test of Cognitive Ability, Stanford-Binet, as well as others (Matthews & Lassiter 2007). McKelvie (1989) determined that the WPT was reliable and useful as a measure of general intelligence in his work with college undergraduates (pp. 161-2). In a report for the Wonderlic organization, it was noted that the WPT, like any reliable measure of cognitive ability is a good predictor of general job performance and training success. This report involved a review of studies on the WPT and prompted the conclusion that it should deliver a higher validity for predicting job performance than other procedures and is justified scientifically and practically as a measure of cognitive ability and indicator of subsequent trainability and job performance (Hunter, 1989). Since teaching involves continual learning and self-direction on the part of the teacher, it is reasonable to assume the WPT would also be a reliable predictor of that component of teaching. The teacher has positional authority as a leader both in the classroom and in the school as a professional learning community. It is appropriate that self-development as a leader, such as required by teachers is affected by the mental ability of the individual and those with higher levels of intelligence should have more capacity to develop themselves as a teacher (Boyce 2004, p.8).The choice of the WPT as the instrument for cognitive ability of the participants in this study was based on the previous observations of consistent reliability, documented validity, acceptable cost, and ease of administration. The WPT will give a concise numerical score that is not subject to misinterpretation which will allow comparison with several different groups of teachers taking the test at different times and locations.

Measuring Subject Matter Knowledge of Teachers

The issue of having effective teachers in our nation's classrooms often stirs much debate over the role that content knowledge plays in teacher effectiveness. NCLB placed the major emphasis on content knowledge to determine if a teacher was highly qualified and most states tie content knowledge into certification requirements as well. Along with the typical requirements for completion of an appropriate post-secondary teacher education program, the use of standardized tests to determine content-knowledge level is often required prior to issue of teacher certification.

Goodman, Arbona and De Rameriz (2008) examined the issue of content knowledge testing by teacher candidates and expressed that there should be some concern about the exclusion of some teacher candidates that may be otherwise capable of being effective teachers based upon these test scores. On the other hand, some states grant certification to teacher candidates with a college degree if they can attain a passing score even if college coursework involved no or very little pedagogy or evaluation of teaching skills. They expressed the need for examination of validity of these tests by using more authentic measures of teaching effectiveness, especially the relationship of these measures to student achievement (p. 24).

Educational Testing Service (ETS) is widely used by state certification agencies through the use of their Praxis™ test series. Forty-five states and the District of Columbia use the Praxis™ tests as part of their certification process. The assessments commonly used are the Praxis I™ which measures basic academic skills, and the Praxis II™ which is designed to measure the teaching candidate's content knowledge for their specific area along with common teaching skills (Praxis™ series information bulletin, 2009 p. 6).

But the question remains whether these exams have any predictive ability when it comes to selecting highly effective teachers. This passage sums up much of what has been found through the review of literature concerning teacher testing by states for certification purposes.

> As to certification exams, there is good evidence that teachers' scores on the exams have a modest positive effect on their students' achievement, with the best evidence of an effect in math. But without evidence on the supply effects of certification exams, the net effect remains in doubt. In the area of teacher supply, there is modest evidence that teacher certification requirements shrink the pool of people who pursue teaching careers but virtually no evidence on whether shrinking the pool has had a meaningful effect on student outcomes. And finally, in the area of hiring, the evidence suggests that schools have limited ability to identify in prospective teachers the attributes that allow them to improve student outcomes (Boyd, et al., 2007).

Since Kentucky requires teacher candidates to take these tests to gain certification as an elementary teacher, use of scores on the Praxis II Elementary Education: Content Knowledge Test Code: 0011 was chosen to quantify this study's fourth and fifth grade teachers' level of content knowledge. Use of that test in Kentucky began September 1, 2005 and was preceded by Praxis II Elementary Education: Curriculum, Instruction, and Assessement Test Code: 0014. Scores on these two required tests were collected to compare with the MAP growth index scores to indicate student achievement in the selected teacher's classroom.

Measuring Teaching Dispositions

Through direct evidence as well as observation, educators know that highly effective teachers have skills that differ from their content knowledge. Having high levels of necessary teaching abilities in addition to content knowledge should make a measurable difference in student achievement in the classroom. We have previously discussed research showing the effects on student achievement from multiple years with highly effective or ineffective teachers which indicate that "the schools students attend and what their teachers know and do are more important influences on student achievement than family characteristics and ethnicity" (Kaplan & Owings, 2001, p. 687).

Acknowledging that great teachers have something more than their intellectual knowledge and cognitive ability working for them in the classroom, is the next step in deciding what to look for in potential teacher candidates. As noted by Liesveld and Miller (2005), great teachers are unorthodox in the way they operate, depending on their own instincts and developing strategies that are successful in reaching students to attain their own educational goals. Research spanning more than forty years indicates that what sets these great teachers apart is individual talent (p.15-16).

One of the factors that can have a direct effect on student achievement is the personality traits or dispositions of the teacher. Singh et al. (2007) noted that in 2006 NCATE defined dispositions as "values, commitments, and professional ethics that influence a teacher's behavior toward his/her students, families, colleagues, and communities" (p.4). They designed a study to measure teacher candidates' dispositions by creating a self-assessment instrument they called the Eastern Teacher Dispositions Index which dealt with five types of personal perceptions. These five areas are:

> (1) Perceptions about self, (2) Perceptions about others, (3) Perceptions about subject field, (4) Perceptions about the purpose of education and the process of education and (5) General frame of reference perceptions (p.8).

In their work, they conceded that there are many ways to define dispositions without general agreement of researchers as to the exact character of teacher dispositions and no universal method to directly measure teacher dispositions. However, the use of surveys, interviews, assessments, or direct observation can be employed to quantify teacher dispositions (p.17).

Thornton (2006) cites several researchers in relating recent work on dispositions to the fairly new concept of emotional intelligence and work by Ritchhart in 2001 to establish a general understanding of dispositions. She wrote:

> Building on Dewey's (1922,1933) work, which addresses the cultivation of habits of mind necessary to effective teaching, Ritchhart (2001) views dispositions as a collection of cognitive tendencies that capture one's patterns of thinking. Ritchhart's definition is grounded in a dispositional view of intelligence and is premised on the concept that "intelligent performance is more than an exercise of ability.... dispositions concern not only what one can do, one's abilities, but also what one is disposed to do. Thus dispositions address the often-noticed gap between our abilities and our actions" (Thornton, 2006).

Justice & Espinoza, (2007) stated that Goleman's Emotional Intelligence Scale gives us an assessment of a person's measure of ten emotional skills. They classified these skills in four performance areas of life: (a) Interpersonal (b) Communication under stress, (c) Personal leadership, (d) Self management in life and career, and (d) Intrapersonal

development (p. 456+). These are very comparable to the five personal perceptions of Singh et al. (2007) and allow us to connect our dispositions discussion to the recent work on emotional intelligence and the role these play in teacher effectiveness.

Because it is so critical to the children in our public schools, we cannot afford to disregard methods of predicting future performance by the teachers we hire. One of the ways school district personnel can predict success is through establishment of a candidate's dispositions toward teaching. These dispositions have been defined as teaching strengths, which, as defined by Gordon (2006), exist when knowledge and skills are combined with talent.

Two on-line assessments, the TeacherInsight™ from Gallup Inc. and Star Teacher PreScreener (STP) from the Haberman Educational Foundation are available commercially to school districts to assist in predicting teacher success. The TeacherInsight ™ replaced the older Teacher Perceiver Interview to measure teaching talent and indicate the level of teaching strengths that a person holds (Gordon, 2006, p. 118). Similarly, the STP evolved from the Urban Teacher Selection Interview (UTSI), which was developed by Marvin Haberman after many years of research involving teachers in urban schools in high poverty settings. (Haberman, 1995).

A written version of the STP was developed around 1990, followed by the on-line version selected for this study. The STP was not designed as a substitute for the oral UTSI, but should screen out applicants that have a low probability of passing the UTSI. The written pre-screener utilizes 50 questions to examine the seven teaching attributes that are evaluated in the UTSI along with three additional functions.

The original seven functions from the interview as represented on the pre-screener

include:

1. Persistence
2. Values children's learning
3. Theory to practice
4. Approach to "At-riskedness"
5. Approach to students
6. Approach to bureaucracy
7. Fallibility

The pre-screener also examines three additional functions that the in person interview does not reflect. The three additional functions are:

1. Planning and organization
2. Explaning teacher success
3. Explaining student success. (Price-Pillow, 2003, pp. 33-34).

The Haberman Educational Foundation offers the STP to school districts as an on-line selection instrument to efficiently pre-screen teacher applicants to become candidates for personal interviews. The STP also has fifty questions dealing with the same areas of teacher qualities found in the UTSI and the three additional areas that Haberman has declared predictive of successful teachers in urban, high-poverty schools (Pillow-Price, 2003).

It is reasonable to assume that the teacher attributes Haberman (1995) identified as effective for teachers that work with urban children living in poverty would also be effective in other situations, but very little research has been completed in this area. Price-Pillow (2003) indicated that there was a shortage of research on effective teacher

attributes in rural schools due to most studies concentrating on teacher recruitment and retention in rural settings with high poverty rates rather than on selection processes. However, rural districts could benefit greatly from a screening process that would predict student success as well as longevity in the teaching career.

Most districts have an immense need to improve the screening process to identify and hire the most effective teachers available. Currently, most school districts use some type of interview as the mainstay of their screening before hiring and more progressive schools may use personnel assessments to give a predictor of the type teacher the candidate might become. We need more information on how these assessments are used to select teachers and whether they are effective in identifying the best teachers from a pool of candidates. More evidence on this issue could help school systems greatly (Jacob, 2007).

Summary

The importance of hiring and retaining very effective, high quality teachers has been emphasized in this review of literature as to the huge impact these decisions can have upon school districts, individual schools, and most importantly upon students. The quality of selection processes varies widely among public schools in Kentucky as well as nationally depending on how much the school district administration has embraced recent research on teacher impact on student achievement.

Even when the district emphasizes a need for hiring the very best teachers available, the modification of selection processes may be hindered by a lack of proven screening tools and appropriate practices that will assist in predicting teacher effectiveness. By determining whether a relationship between student achievement and

teacher test scores exists will help districts develop screening procedures that include useful parameters in the teacher selection process. This research project will add to the information available to educators when deciding where to invest time and money in development of school districts' teacher selection process.

CHAPTER III

METHODOLOGY

Identification of the Research Problems

The purpose of American public schools is to produce student learning that will prepare them for a successful life. There has been much work done to establish national and state standards for student learning with a large amount of effort devoted to measuring student achievement in our public schools. Many variables are involved in student learning rising both from in-school and out-of-school issues. In many instances, school personnel have little control over these variables but attempt to do all they can to advance student achievement. They do have control over an obvious variable, the effectiveness of the classroom teacher.

Many requirements are made on prospective teachers, first by universities, then by States, and finally by the school districts. Little research has been done to verify that the resulting data are fully utilized by school districts that are trying to select the best teachers available. The purpose of this research is to investigate the possible relationships of three common attributes in teacher selection processes; teacher dispositions, teacher intelligence, and teacher content knowledge, with academic achievement of students in the teacher's classroom. The researcher hopes that the proposed study will add to scholarly research that school human resource professionals may use when developing policy and procedures to use in hiring the best teachers available.

This researcher selected three attributes of teachers that are commonly used to

measure effective teachers, intellectual ability, content knowledge, and teaching dispositions to compare to student achievement in their classrooms. These three teacher attributes and classroom student achievement are easily quantifiable using assessments that are widely used today. Three research questions directly related will be used to direct this research, they are:

Research Question 1: Is there a relationship between the academic progress of students in reading and math and the cognitive ability of their teacher as measured by the Wonderlic Personnel Test?

Research Question 2: Is there a relationship between the academic progress of students in reading and math and the rating of teacher attributes as measured by the Star Teacher Pre-Screener™ on-line assessment?

Research Question 3: Is there a relationship between the academic progress of students in reading and math and the teacher's knowledge of pedagogy/subject matter as measured by the Praxis II™ test for elementary teachers?

Research Design

This project is a quantitative study to identify possible relationships between three common assessments that are being used by school districts in pre-hire evaluation of teacher candidates and student achievement in classes taught by those teachers. Cognitive ability of the teacher will be measured by the Wonderlic Personnel Test, teaching strengths will be measured by the Star Teacher Pre-Screener assessment, and subject matter knowledge will be measured by the PRAXIS II test required for elementary teacher certification in Kentucky. Student achievement will be measured with the Growth Index Scores on the MAP assessment over two school years, 2007-08 and 2008-09. More

explanation of these tests follows in the assessments section of this chapter.

Study Population

This researcher sought to include only fourth and fifth grade teachers and classrooms in this study for similar reasons to those considered by William Sanders and his associates for research in Tennessee that resulted in very rich conclusions about the effect that teachers have on individual student achievement. Most elementary teachers at this level teach all subjects to a self-contained classroom or specific subjects to students within a small team of teachers and students (Wright et al., 1997)

In order to limit outside influences on student achievement data, sample integrity demands that the teachers included in this study have the major responsibility for teaching reading and/or math at the fourth and fifth grade levels. Certain teachers within schools may not qualify for inclusion in this study if they did not teach reading and/or math to fourth and/or fifth grade students in a whole class setting. Reading and math MAP test rosters will be produced listing student achievement scores for students in those classrooms.

Four selected Kentucky school systems were asked to be a part of the study. School Superintendents in these four county school districts were invited to participate in this research project. They were chosen due to use of MAP testing for the years involved in this research along with their use of the TeacherInsight™ assessment in their teacher selection process. The first step requires establishment of their willingness to allow the research in their schools.

The school districts that respond to the request to conduct this study will facilitate contact with the principals of elementary schools that house fourth and fifth grades.

Teachers within the participating schools must volunteer to become a part of this study and agree to be tested using the WPT and the on-line STP. An incentive for volunteering to be included the study will be provided through random drawings of participant names to receive twenty dollar gift cards.

The rate of participation can vary widely across the districts, but the eligible teacher population of eligible fourth and fifth grade teachers could be substantial. Some of the teachers may have taught one students in one grade, while others could teach both levels. There are teachers that teach both reading and math while others teach single subjects to multiple classrooms through teaming arrangements with other teachers. To a limited extent, teachers might have taught reading one year and math the other year.

Data Analysis

Analysis of the data will consist of linear regression modeling based upon the Pearson Product Moment correlation to determine the strength of a linear relationship between the teacher assessment variable and the two student achievement variables included in each of the three research questions. Further observational study of the data may be considered as needed for use of different comparison models or the analysis of correlation through multiple regression to further evaluate the teacher quality variables; cognitive ability, content knowledge, and teaching disposition in regard to student achievement.

Pilot study

A pilot study will be conducted to determine the effectiveness of the planned protocol and preparation for the completion of the larger research project that involved larger Kentucky school districts. The pilot school district has one elementary school

housing fourth and fifth grades. Both fourth and fifth students are assigned to teams of teachers for two years, an arrangement commonly referred to as looping. This arrangement will allow the use of both fourth and fifth grade student achievement data for most of the participants.

Following a meeting with the superintendent and the principal of the pilot study district, the researcher will attend a faculty meeting to outline the research project and enlist volunteers for the pilot study. Many of these teachers have already taken some of the required assessments, but a few may have to complete the WPT to be included in the pilot study. The teachers will have one week to decide to participate. All teachers wishing to participate will be asked to sign a subject informed consent form and a short informational sheet to allow their individual data to be correctly entered into the study.

When the teachers decide to join the study, an identification code number will be assigned to replace their name in reference to all data. These code numbers with the corresponding individual names will be located on the informational sheet and stored separately from the teacher and student test data during the course of the study. The Praxis II, WPT, and STP scores will be recorded in an excel spreadsheet to facilitate analysis using the code numbers.

Assessment Data

The teachers will be asked to complete the STP on-line assessment at their leisure using instructions provided electronically from the researcher. The results of this assessment are immediately available to the researcher through an account established with the Haberman Educational Institute. The STP reports both a numerical score comprised of the correct responses from the fifty multiple choice questions on this

assessment as well as a report of ratings for the ten attribute areas. Only the numerical score is considered in the quantitative analysis of this study.

Praxis II test scores are available for a portion of the participants in the study but may be limited due to the fact that these test scores are not available for veteran teachers. Teacher assessment scores on Praxis II will be self-provided or obtained from existing data in the teacher's records. These records are traditionally kept at the district office and at the Kentucky Education Professional Standards Board for all teachers as part of the certification process.

The student MAP scores will be aggregated at the classroom level by using the growth index (GI) score. The individual student data will be analyzed to identify and remove invalid scores that exceed three standard deviations from the mean score for the classroom. The next step will be to aggregate the individual student GI scores to determine the average academic growth of that class of students. Classroom GI scores for teachers with more than one class in the same subject area will be combined for an average GI. Next, the average GI data for each teacher will be combined for both grades over the two years. Final GI data for each teacher shall be analyzed by subject area to determine any possible relationships that might exist.

Delimitations

Several delimitations may affect the results of the study, including:

- The participants in the study are limited to schools where the principals facilitate access to teachers by the researcher and where the district will supply student achievement data.
- Only teachers that volunteer are included in this study resulting in a non-random

population of cooperative teachers. The teacher population may lack the diversity to establish a widely applicable conclusion.

- This study includes existing scores for the Praxis II test, code number 0011 and 0014 and handled separately giving a very small sample size for each test. Test data from both versions of the test includes only the less experienced teachers and does not exist for the veteran teachers included in this study.
- The on-line STP is not a proctored assessment and an assumption is made by the researcher that the teacher will honestly respond to the questions without any influence by another person or situation.
- The study uses one measure of student achievement; the MAP Growth index average for reading and math classrooms. This instrument is limited to multiple choice selections on a computer and does not include other measures of student learning such as performance events such as open response items in reading and math.
- The experience level of the teachers included in this study is not differentiated for comparison. Literature review indicates that teaching experience may affect the effectiveness of a new teacher but is not a significant factor in teacher effectiveness after five years of experience.
- Because the STP results will be on teachers that are currently employed for three or more years in the school district providing student and teacher data, no conclusion may be drawn concerning its effectiveness for hiring new teachers.
- Many variables in the classroom and outside of school affect student learning but are not considered in this study. The preponderance of research on this subject

indicates that teacher effect on student learning is not dependent on ethnicity, academic achievement level, or socio-economic status of the students. Therefore, these potential variables are noted but not dealt with directly in this study.

CHAPTER FOUR

RESULTS

Introduction

The purpose of this study was to determine if any relationship could be found between student achievement in the classroom setting and the test scores of current teachers on three common pre-employment tests. The teachers that volunteered for this study provided test results from the Wonderlic Personnel Test (WPT), Praxis II test for elementary teachers, and the Star Teacher Pre-Screener (STP) assessment. All participants were given the WPT and STP assessment by the researcher, while Praxis II data was gathered only on those teachers that had taken that test as a necessary step in the Kentucky teacher certification process. Participating teachers taught reading and/or math at the fourth or fifth grade level during the 2007-08 and/or the 2008-09 school years. Selected school districts supplied student achievement data for classrooms of the participating teachers using data from the Measures of Academic Progress (MAP) computerized adaptive test. Appendix A lists all teacher test score data from the sample population with corresponding classroom GI scores for math and reading.

Results from statistical data analysis are presented for these three research questions:

1. Is there a relationship between the academic progress of students in reading and math and the cognitive ability of their teacher as measured by the Wonderlic Personnel Test™?

2. Is there a relationship between the academic progress of students in reading and math and the rating of teacher attributes as measured by the Star Teacher Pre-Screener™ on-line assessment?
3. Is there a relationship between the academic progress of students in reading and math and the teacher's knowledge of pedagogy/subject matter as measured by the Praxis II™ test for elementary teachers?

Three research hypotheses guided the analysis of data, with each null hypothesis tested at the $p < .01$ level:

Null hypothesis 1: Scores on the Wonderlic Personnel Test have no relationship to the corresponding academic progress of students in the classroom.

Null hypothesis 2: Scores on the Star Teacher Pre-Screener™ have no relationship to corresponding academic progress of students in the classroom.

Null hypothesis 3: Scores on the Praxis II™ have no relationship to the corresponding academic progress of students in the classroom.

Pilot Study

The Pilot Study proved the research protocol to be workable with one refinement. It was not feasible to attend a faculty meeting at all the schools in the larger districts. Unlike the pilot district, most public school districts' elementary schools house all grades from pre-school through fifth or sixth grades resulting in fewer fourth and fifth grade teachers in each building. In these districts, principals facilitated meetings with eligible participants due to the limited number of teachers in each school that could provide appropriate data. Time was the biggest restraint on this research project and became apparent through the pilot study.

There were nine teachers in the pilot study comprised of eight female and one male. Five teachers had taken the Praxis II Code 0011 test which indicated that they had taken the test before 2005 and are identified in this research study as mid-veteran teachers. Four teachers took the Praxis II Code 0014 test, instituted as a certification requirement for Elementary teachers beginning in September of 2005. These individuals are identified as non-veteran teachers in this research study. Non-veteran teachers had no more than three years of experience during the time period that MAP GI scores were considered for the study. There was no racial diversity and practically no gender diversity within the pilot study group.

All of the pilot study teachers completed the STP assessment during the study. WPT scores for five of the teachers were obtained from personnel records and four mid-veteran teachers took the WPT during the pilot study. All teachers self-reported their Praxis II scores to the researcher. Student achievement for the pilot study was analyzed using the average GI of each teacher in the appropriate subject area, reading or math. The pilot study data analysis showed no significant relationship between classroom student achievement and any of the teacher test scores. It must be noted that the population of teachers in the pilot was extremely small and that the pilot study population was included in the full research study.

Full Research Study Data Analysis

Three of the four districts invited to participate were included with varying rates of participation. One district did not respond to the researcher requests and no teacher or student data was available. The school districts that responded positively to the research request allowed the researcher to contact elementary school principals of schools that

house fourth and fifth grades and were invited to participate. Due to year to year mobility of teachers within the school or district, a portion of the teachers had data for only one year while others provided data for two consecutive years. Some of the teachers taught only students in one grade, while others could teach both levels. In all cases, student achievement data for the appropriate teacher and subject area was carefully matched. This is reflected in the variability of sample numbers reported for each subject area from two years of student data.

The goal of a diverse teacher sample was not reached due to the self-limiting nature of voluntary participation by teachers. All teachers included in this study were white and only one male teacher was included in the pilot study. Although qualitative data was not gathered as part of this study, it is important to note that most of the teachers participating were veteran teachers, with only three teachers providing student data from their first year of teaching. This preponderance of veteran teachers in the sample resulted in a very low number of participants with Praxis II scores and ultimately limited the reliability of that portion of the study. These elements are noted as limitations later in this chapter.

There were thirty-two teachers that volunteered and met the qualifications to be in the research study. The research study population comprised of thirty-one female and one male. One teacher withdrew from the study before completion of the WPT or STP. One teacher's data was not included in data analysis in order to maintain the integrity of the data. This exclusion was necessary due to a very small student population in only one subject area for only one year. The remaining population of thirty white, non-Hispanic teachers exhibited no racial diversity and included only one male.

The teachers included in the data analysis completed the WPT and STP assessments and had student achievement data for math and/or reading during at least one of two school years, 2007-08 and/or 2008-09. Six non-veteran teachers took the Praxis II test code 0011, and six mid-veteran teachers took the Praxis II test code 0014. Teachers were allowed to self-report Praxis II scores that were missing from district personnel records and are included in this study. The teacher scores from the two versions of the test were not co-mingled when correlation computations for this component were conducted.

Participants completed the WPT which was administered to the teachers by this researcher at the schools where they work. The researcher proctored the timed pencil and paper WPT with the time and location determined by the school principal and the teachers involved. The score on the WPT is reported as the number of correct answers from the fifty questions on this test during the thirty minute testing period.

The teachers were also asked to complete the STP on-line assessment at their leisure using instructions provided electronically from the researcher. The results of this assessment are not known to the participant but are immediately made available to the researcher through an account established with the Haberman Educational Institute. The STP reports both a numerical score comprised of the correct responses from the fifty multiple choice questions on this assessment as well as a report of ratings for the ten attribute areas. Only the numerical score is considered in the quantitative analysis of this study.

As noted previously and later in the limitations, the number of Praxis II test scores available for the study were very low due to the number of veteran teachers in the study. Also limiting this portion of the study, scores on the latest version of the Praxis II test

(0014 Elementary Content Knowledge) are not comparable to the former test (0011 Elementary Curriculum, Instruction, and Assessment) required for elementary teachers and these test scores are not available for veteran teachers. Teachers were allowed to self-report Praxis II scores because many of reports were missing from district personnel record.

Student Achievement Data

The student MAP scores were aggregated at the classroom level by using the growth index (GI) score. The individual student data was analyzed to identify and remove invalid scores which exceeded three standard deviations from the mean score for the classroom. Outlier scores could have resulted in student apathy or an unusual testing situation during either the fall or spring MAP testing period and as expected, very few outlying student scores had to be excluded.

The next step involved aggregating the individual student GI scores to determine the average academic growth of that class of students. Classroom GI scores for teachers with more than one class in the same subject area were combined for an average GI. Average GI data for each teacher was analyzed by averaging GI data for both grades over the two years for teachers teaching both grade levels in the same subject area. Final GI data for each teacher was analyzed by subject area (Appendix B).

Teacher Cognitive Ability

General cognitive ability of the teachers is represented by scores on the WPT which includes fifty questions with the score representing the number of correct responses. The WPT scores had a range of 16 points, with a low score of 25 and high score of 41. The mean (33.9) and median (34) indicates a tight sample for the group with

no outlying scores. Data analysis showed no significant relationship between the WPT test scores and the classroom Math GI (Table 4.1) or Reading GI (Table 4.2). Analysis of correlation was performed using Pearson Product Moment correlation one-tailed test at the $p < .05$ significance level. The conclusion can be drawn that there was no significant relationship between classroom student achievement and teachers' WPT scores.

Null hypothesis 1 is proven. Scores on the Wonderlic Personnel Test have no significant relationship to the corresponding academic progress of students in the classroom in the sample population.

Table 4.1

Pearson Correlation of WPT to Math GI

		WPT	MathGI
WPT	Pearson Correlation	1	-.285
	Sig. (1-tailed)		.099
	N	30	22
MathGI	Pearson Correlation	-.285	1
	Sig. (1-tailed)	.099	
	N	22	22

Table 4.2

Pearson Correlation of WPT to Reading GI

		WPT	ReadingGI
WPT	Pearson Correlation	1	-.371
	Sig. (1-tailed)		.041
	N	30	23
ReadingGI	Pearson Correlation	-.371	1
	Sig. (1-tailed)	.041	
	N	23	23

Rating of Teacher Dispositions

The STP used to derive a rating of teacher dispositions includes fifty questions with the score representing the number of correct responses. The STP scores had a range of 14 points, with a low score of 26 and high score of 40. The mean (33.7) and median (34) indicates a tight sample for the group with no outlying scores. The Pearson Product Moment correlation test showed no significant relationship between the STP test scores and the classroom Math GI (Table 4.3) or Reading GI (Table 4.4). Analysis of correlation was performed using Pearson Product Moment correlation one-tailed test at the $p < .05$ significance level. One can infer from the data analysis that no significant relationship can be found between classroom student achievement and teachers' STP scores.

Null hypothesis 2 is proven. Scores on the Star Teacher Pre-Screener have no relationship to the corresponding academic progress of students in the classroom in the sample population.

Table 4.3

Pearson Correlation of STP to Math GI

		STP	MathGI
STP	Pearson Correlation	1	.020
	Sig. (1-tailed)		.465
	N	30	22
MathGI	Pearson Correlation	.020	1
	Sig. (1-tailed)	.465	
	N	22	22

Table 4.4

Pearson Correlation of STP to Reading GI

		STP	ReadingGI
STP	Pearson Correlation	1	-.245
	Sig. (1-tailed)		.130
	N	30	23
ReadingGI	Pearson Correlation	-.245	1
	Sig. (1-tailed)	.130	
	N	23	23

Subject Knowledge of Teachers

This portion of the research project is probably the least predictive of the three areas examined due to the very small sample size and problems associated with changes in the Praxis II tests required by Kentucky for teacher certification. Therefore, the total population (n=12) of teachers providing Praxis II scores was very low even before splitting the sample for analysis.

The population of non-veteran teachers that took the Praxis II test code 0011 was equal (n=6) to the population of mid-veteran teachers (n=6) that took the Praxis II test code 0014. Since the two versions of the Praxis II are not comparable due to different content and test construct, the two versions of the test were not co-mingled for comparison. Data analysis indicates that there was no significant relationship between classroom student achievement and any of the teacher test scores as neither sub-group showed any significant correlation to the classroom GI math scores (Table 4.5) or reading scores (Table 4.6), when an analysis of correlation was performed using Pearson Product Moment correlation two tailed test with a significance level of $p < .05$.

Null hypothesis 3 is proven. Scores on the Praxis II tests have no relationship to the corresponding academic progress of students in the classroom in the sample population.

Appendix B shows all correlation figures for the research study, the only report of significance was minor and was not directly related to the stated research questions. For example, there was a weak Pearson correlation (.780 $p < .01$) between Math GI and Reading GI for teachers (n=15) teaching both subjects suggesting that teachers with successful students in one subject had successful students in the other subject as well. It should be noted that a very weak negative correlation (-.371 $p < .05$) is reported in the Pearson correlation table (Appendix B) between WPT scores and Reading GI but is not indicative of any true relationship.

After examination of the Pearson Correlation data analysis, the researcher decided that further analysis of the data was not indicated with the Praxis II data because of the small sample size and no significant correlation between either Praxis II test scores and

student achievement in MathGI or ReadingGI could be determined at the significance level of $p < .05$.

Table 4.5

Pearson Correlation of Praxis II tests to Math GI

		Praxis0011	Praxis0014	MathGI
Praxis0011	Pearson Correlation	1	.[a]	.833
	Sig. (1-tailed)		.	.187
	N	6	0	3
Praxis0014	Pearson Correlation	.[a]	1	-.379
	Sig. (1-tailed)	.		.376
	N	0	6	3
MathGI	Pearson Correlation	.833	-.379	1
	Sig. (1-tailed)	.187	.376	
	N	3	3	22

Table 4.6

Pearson Correlation of Praxis II tests to Reading GI

Praxis0011	Pearson Correlation	1	.[a]	.244
	Sig. (1-tailed)		.	.378
	N	6	0	4
Praxis0014	Pearson Correlation	.[a]	1	-.471
	Sig. (1-tailed)	.		.212
	N	0	6	5
ReadingGI	Pearson Correlation	.244	-.471	1
	Sig. (1-tailed)	.378	.212	
	N	4	5	23

a. Cannot be computed because at least one of the variables is constant.

Regression Analyses

The reseacher performed two additional analyses to determine if a combination of variables measured on teachers might have any significant relationship to the student educational outcomes *MathGI* score and *ReadingGI* score. Two multiple regression analyses were performed. In the first, the predictor variables were STP score and WPT score and the dependent variable was MathGI score. In the second regression, the same predictors were used, with the dependent variable ReadingGI score. In each regression analysis, the two predictors were entered into the equation simultaneously.

Results of the first regression were that there was no significant relationship between the predictors' STP score and WPT score and the dependent variable MathGI score, $F(2,19) = 1.09$, $p = .36$. The percentage of variance accounted for by the predictors was $R^2 = .103$, (adjusted $R^2 = .009$). Thus, the variance in the dependent variable MathGI score accounted for by the two predictors was very small and not statistically significant.

Results of the second regression were similar to the first analysis. There was no significant relationship between the predictors STP score and WPT score and the dependent variable ReadingGI score, $F(2,20) = 1.64$, $p = .22$. The percentage of variance accounted for by the predictors was $R^2 = .141$, (adjusted $R^2 = .056$). The variance in the dependent variable ReadingGI score was small and not statistically significant.

CHAPTER FIVE

DISCUSSION AND CONCLUSIONS

Introduction

When considering the time and effort put into teacher hiring by school districts and the lack of uniformity of the hiring process between school districts, the decision about what is the most effective direction to pursue is nearly impossible to make with any degree of confidence. Through a review of current literature and research studies showing the lasting impact that teacher quality has on student success (Sanders & Rivers, 1996), it is obvious that astute school leaders should place emphasis on teacher quality when deciding where to spend limited school funding. Devoting time and resources to improve the processes that are used to screen and hire teachers is necessary and feasible and should be continued. The cost of teacher turnover along with the potential negative impact on individual student learning should encourage educational agencies to continue to evaluate teacher selection tools and processes that may effectively predict teacher performance in the classroom.

Although, the results of this small study do not indicate that the three teacher selection tools can be significantly correlated to student achievement, school leaders should be cautioned not to take this as a reason to devote less effort in that endeavor (Lee, 2005). This challenge of selecting effective teachers from hosts of applicants can be daunting, demanding use of tools and procedures that will reliably predict teachers' ability to combine multiple qualities to positively affect their students. As pointed out by

Kirchner (2008) “effective teaching requires a combination of cognitive, behavioral, and dispositional traits” (p 136). For economic and societal reasons pointed out earlier, school districts cannot wait for new teachers to prove themselves on the job.

Discussion and Implications

The study was limited due to lack of participation by teachers (N= 30), specifically in the largest school district included in the study. Unfortunately, this reduced the final population significantly from what was proposed at the beginning of this research study. It would not be prudent for readers to infer generalizations from this study, however the results should assist in the development of future research projects that need to be completed.

Some of the issues that limited this study could be avoided in future studies of this nature and need to be discussed further. First, if this study had been conducted early in the school year with most of the data gathering done in the fall semester, more effort could have been given to recruitment of teacher participants. The pilot study had 90% (nine out of ten) of the eligible teachers to participate compared to less than 20 % (seven out of forty) in the last district admitted to the study. The effort to collect data well into the spring semester was complicated by several negative conditions. Bad weather caused all the participating districts to miss several days of school which made it difficult to keep on a strict timeline. Teachers and principals generally are placed under increasing pressure from many directions as school year progresses.

Secondly, the fact that the majority of schools that participated in this study were from rural areas should be considered in reviewing the results. Diverse teacher and student populations are not included in this study and the results may only be indicative

of rural elementary schools working with a low level of students living in poverty. Haberman (2004) points out how the intricacies of working within the educational construct of the classroom and general culture of the classroom can profoundly affect the success of teaching strategies. It is possible that the teaching dispositions required for teachers in rural schools could be very different from those of an urban district with a high percentage of students living in urban poverty. This observation is similar to that concluded by Klussmann (2004), where student achievement on the Texas Assessment of Academic Skills was used to examine possible relationships between the student data of teachers selected using traditional teacher selection to corresponding data from teachers selected using Haberman's Star Teacher selection methodology.

Third, examination of the teacher sample (appendix A) shows that 60% of the teachers did not have Praxis II scores placing them in the veteran teacher category. Teacher turnover is most common in the non-veteran population of school districts for many reasons, including non-rehiring and self-selection of many ineffective teachers' (Skandera & Sousa, 2000). This implies that the study sample would hold a higher percentage of effective teachers that have gained many skills that might not be reflected in the teacher selection tools evaluated in this study.

Fourth, this researcher observed that a cooperative administrative culture was obvious in those schools where a large percentage of the teachers participated in the study. Because of the protocol followed in selecting teacher participants, a small sample with very little variability among the teacher population was obtained. Anecdotal evidence from this researcher's experience during the study indicates a higher percentage of teachers participated in districts where the superintendent communicated support for

the research study to the principals and teachers.

Shutt (2004) found that school culture significantly impacted student achievement in that school. It is possible that the school culture promulgated by a principal may profoundly affect the teachers' performance in the classroom too. Anecdotal evidence from the schools in this study and the specifically in the pilot study indicated that an unstable situation occurred during the years that student data was gathered involving principal turnover and a complete change in the school configuration during the year prior to data collection. It could be that this contributed to lower MAP GI scores compared to very stable schools included in the study.

Lastly, the student assessment process using MAP is highly structured and includes parameters to allow uniform reporting of GI across a wide range of student ability levels (Cronin & Bowe, 2006), however it must be noted that individual schools have varying situations as to when, where, and how the testing takes place that may affect student performance. It is possible that comparison of scores from one district to another could inherently carry some variables that are not examined in this research study. Although, not reportable in this study due to the need to protect against unintended identification of participants, there seemed to be a definite difference in the range of student GI scores when comparing from one district to another.

Suggestions for Future Research

Looking at the results from a purely quantitative standpoint we can state that these three teacher selection tools do not show any significant correlation to the student academic achievement for the study participants over the two year period 2007-2009. The limitations of this study should result in an end to this line of research but rather should

stimulate a more expansive drive to answer questions. What can be learned from this study is how to design and carry out a more powerful project that would overcome the noted limitations in this study. This might allow wider application of the findings.

Several things could be considered to further study in this area and are listed here.

- The same study might be replicated in a large urban setting with a more diverse teacher sample. This would result in more reliability of the results by increasing the range of scores on the teacher instruments and having a more uniform reporting of student achievement data.
- The study could be replicated across many Kentucky districts to include only non-veteran teachers in the sample to include more teachers that might be classified as ineffective. This could involve testing of non-veteran teachers early in the school year and gathering student achievement data for their first two years of teaching.
- Redesigning the study to use an instrument that purports to measure different teaching dispositions that may be more applicable to the general student population. STP was specifically designed to predict success when working with at-risk students.
- A meta-analysis of similar studies might be useful to point researchers in new directions in an effort to find ways to predict teacher effectiveness.

Further study using different teacher selection tools, data gathering processes, or student achievement measures could give insight into determining particular teacher attributes that could be measured and declared predictive of advanced student performance. The need is too great and the challenge too important to stop looking.

REFERENCES

Amrein-Beardsley, A. (2008). Methodological concerns about the education value-added assessment system. *Educational Researcher, 37*(2), pp. 65-75.

Andrew, M. D., Cobb, C. D., & Giampietro, P. J. (2005). Verbal Ability and Teacher Effectiveness. Journal of Teacher Education, 56(4), 343+.

Arnold, D. (2006). *Cognitive ability testing*. Wonderlic Inc. white paper. Libertyville, IL.

Ascher, C. & Fruchter, N. (2001). Teacher quality and student performance in New York City's low performing schools. *Journal of Education for Students Placed At Risk, 6*(3), 199-214.

Barnes, G., Crowe, E., & Schaefer, B. (2007). *The cost of teacher turnover in five school study.* National Commission on Teaching and America's Future. Washington, D.C.

Blanton, L. P., Sindelar, P. T., & Correa, V. I. (2006). Models and measures of beginning teacher quality. *Journal of Special Education, 40*(2), 115+.

Boyce, L. (2004). *Propensity for self-development of leadership attributes: Understanding, predicting, and supporting performance of leader self-development.* Dissertation article, George Mason University, USA Virginia.

Boyd, D., Goldhaber, D., Lankford, H., & Wyckoff, J. (2007). The Effect of Certification and Preparation on Teacher Quality. *The Future of Children*, 17(1), 45+.

Child Care and Early Learning. (1999, September). *American Journalism Review,* 21.

Cronin, J. & Bowe, B., (2006). *NWEA research report: Study of growth index performance by school.* Lake Oswego, OR: Northwest Evaluation Association.

Eby, F. & Arrowood, C. (1934) The development of modern education: In theory, organization, and practice. New York: Prentice Hall.

The Praxis Series Information Bulletin. (2009). Educational Testing Service. http://www.ets.org/praxis..

Goe, L. (2007). The link between teacher quality and student outcomes: A research synthesis. Washington, DC. National Comprehensive Center for Teacher Quality.

Goe, L & Stickler, L. (2008). Teacher quality and student achievement: Making the most of recent research. *TQ Research and Policy Brief.* Washington, DC. National Comprehensive Center for Teacher Quality.

Goodman, G., Arbona, C., & De Rameriz, R. D. (2008). High-stakes, minimum-competency exams: How competent are they for evaluating teacher competence? *Journal of Teacher Education, 59*(1), 24+.

Gordan, G. (2006). Building engaged schools: Getting the most out of America's classrooms. Omaha, NE: Gallup Press.

Haberman, M. (1995). Selecting 'Star' teachers for children and youth in urban poverty. *Phi Delta Kappan*, 76(10), 777+.

Haberman, M. (2004). Star Teachers. Houston, TX:Haberman Educational Foundation.

Haberman, M. (2007). Learning, growth and socialization. (A paper presented to the Association of Teacher Educators Summer Conference, Milwaukee: July 29, 2007).

Imig, D., & Imig, S. (2006). The teacher effectiveness movement: How 80 years of essentialist control have shaped the teacher education profession. *Journal of Teacher Education, 57*(2), 167+.

Jacob, B. (2007). The challenges of staffing urban schools with effective teachers. *The Future of Children, 17*(1), 129-154.

Justice, M., & Espinoza, Justice, M. & Espinoza, S. (2007). Emotional Intelligence and Beginning Teacher Candidates. Education, 127(4), 456+.

Kaplan, L. & Owings, W. (2001). Teacher quality and student achievement: Recommendations for principals. *NASSP Bulletin, 85*(628), 64-73.

Kirchner, J. (2008). Teacher effectiveness: The relationship of the teacher work sample to the teacher insight interview. Ph.D. Dissertation, University of Louisville, USA – Kentucky.

Klussman, D. (2004). The impact of teachers selected by the Star Teacher Interview process on student achievement. Ed.D. Dissertation, Seton Hall University, USA – New Jersey.

Knudsen, L. & Morrissette, P. (1998). Goals 2000: An analysis and critique. *International Electronic Journal For Leadership in Learning, 2*(4).

Lee, D. M. (2005). Hiring the best teachers: gaining a competitive edge in the teacher recruitment process. *Public Personnel Management, 34*(3), 263+.

Little, Paula S. (1998) The effect of rural values on personnel selection practices in Kentucky public school districts. Ed.D. dissertation, University of Louisville, USA – Kentucky. Retrieved February 6, 2009, from Dissertations & Theses: Full Text database. (Publication No. AAT 9924995).

Liesveld, R. & Miller, J. (2005).*Teach with your strengths*. New York: Gallup Press.

Matthews, T. D. & Lassiter, K. (2007). What does the Wonderlic Personnel Test measure?. *Psychological Reports, 100*(3), 707-712.

McKelvie, S. (1998). The Wonderlic Personnel Test: Reliability and validity in an academic setting. *Psychological Reports 65*(1), 161-162.

Murname, R. & Steele, J. (2007). What is the Problem? The challenge of providing effective teachers for all children. *The Future of Children, 17*(1), 15-43.

Murray, F. B. (2007). Disposition: A Superfluous Construct in Teacher Education. *Journal of Teacher Education*, 58(5), 381.

National Commission on Teaching and America's Future. (2008, April). *Learning Teams: Creating What's Next*. Washington, DC: Author.

Peterson, K. D. (2002). Effective teacher hiring: A guide to getting the best. Alexandria, VA: Association for Supervision and Curriculum Development.

Pillow-Price, K. (2003). Evaluation of the Haberman Urban Teacher Selection interview in rural settings. PhD. Dissertation, Arkansas State University, USA – Arkansas.

Rivkin, S., Hanushek, E., & Kain J. (March 2005). Teachers, schools, and academic achievement. *Econometrica, 73*(2) 417–458.

Sanders, W. & Horn, S. (1998). Research findings from Tennessee value-added assessment system (TVAAS) database: Implications for education evaluation and research. *Journal of Personnel Evaluation in Education. 12*(3), 247-256.

Sanders, W. & Rivers, J. (1996). *Cumulative and residual effects of teachers on future student academic achievement.* Knoxville: University of Tennessee Value-Added Research and Assessment Center.

Scherer, M. (Ed.). (2003). Keeping Good Teachers. Alexandria, VA: Association for Supervision and Curriculum Development.

Shutt, A. (2004). School Culture in Kentucky elementary schools examining the path to proficiency. PhD. Dissertation, University of Louisville, USA – Kentucky.

Skandera, H. & Sousa, R. (2003). *School figures: The data behind the debate*. Stanford, CA: Hoover Institution Press.

Singh, D. & Stoloff, D. (November, 2007). Measuring dispositions. Paper:Presented at the National Fifth Annual Symposium on Educator Dispositions. Erlanger, KY.

Smith, M. L. & Fey, P. (2000). Validity and accountability in high-stakes testing. *Journal of Teacher Education, 51*(5), 334-344.

Stronge, J. H. (2002). *Qualities of effective teachers.* Alexandria, VA: Association for Supervision and Curriculum Development.

Thornton, H. (2006). Dispositions in Action: Do Dispositions Make a Difference in Practice?. *Teacher Education Quarterly, 33*(2), 53.

Tucker, P. D., & Stronge, J. H. (2005). *Linking Teacher Evaluation and Student Learning*. Alexandria, VA: Association for Supervision and Curriculum Development.

Vogler, K. (2002). The impact of high-stakes, state-mandated student performance assessment on teachers' instructional practices. *Education, 123*(1), 39+.

Wright, S. P., Horn S., & Sanders, W. (1997). Teacher and classroom context effects on student achievement: Implications for teacher evaluation. *Journal of Personnel Evaluation in Education. 11*, 57-67.

Yeh, S. (2006). Reforming federal testing policy to support teaching and learning. *Educational Policy, 20*(3), pp. 495-524.

APPENDIX A

Full Study Population Data

	Praxis 0011 Score	Praxis 0014 Score	STP Score	WPT Score	Math GI Score	Reading GI Score
1	165		33	37	-2.87	
2		175	35	33		-1.22
3	181		38	41	-2.40	-3.31
4		164	35	32	-3.81	-3.11
5	187		37	38	-0.54	
6	189		36	39		-2.84
7		192	39	33		-2.24
8	195		40	39		-1.56
9		180	31	34	-1.51	
10			36	36	7.35	4.62
11			40	36		2.46
12			39	35	3.02	-0.05
13			37	32	1.75	
14			31	37		0.97
15		160	27	25	3.25	1.77
16			35	31	6.40	3.26
17			34	27	3.52	2.36
18			37	38	-0.79	3.63
19			34	37	1.64	-0.88
20			34	27	2.73	2.44
21			26	33	3.55	3.20
22			38	27	2.51	
23			32	27	2.48	1.61
24	181		34	35		-1.28
25			35	41	0.81	
26		167	33	26		2.18
27			38	38	6.27	
28			36	33	3.75	4.07
29			32	33	-1.08	-2.56
30			34	32	0.80	1.61
N	6	6	30	30	22	23

APPENDIX B

Pearson Product-Moment Correlations

		STP	WPT	Praxis0011	Praxis0014	MathGI	ReadingGI
STP	Pearson Correlation	1	.402*	.799*	.679	.020	-.245
	Sig. (1-tailed)		.014	.028	.069	.465	.130
	N	30	30	6	6	22	23
WPT	Pearson Correlation	.402*	1	.343	.696	-.285	-.371*
	Sig. (1-tailed)	.014		.253	.062	.099	.041
	N	30	30	6	6	22	23
Praxis0011	Pearson Correlation	.799*	.343	1	.[a]	.833	.244
	Sig. (1-tailed)	.028	.253		.	.187	.378
	N	6	6	6	0	3	4
Praxis0014	Pearson Correlation	.679	.696	.[a]	1	-.379	-.471
	Sig. (1-tailed)	.069	.062	.		.376	.212
	N	6	6	0	6	3	5
MathGI	Pearson Correlation	.020	-.285	.833	-.379	1	.780**
	Sig. (1-tailed)	.465	.099	.187	.376		.000
	N	22	22	3	3	22	15
ReadingGI	Pearson Correlation	-.245	-.371*	.244	-.471	.780**	1
	Sig. (1-tailed)	.130	.041	.378	.212	.000	
	N	23	23	4	5	15	23

*. Correlation is significant at the 0.05 level (1-tailed).

a. Cannot be computed because at least one of the variables is constant.

**. Correlation is significant at the 0.01 level (1-tailed).

CURRICULUM VITAE

NAME: Jon D. Hall

ADDRESS: 1653 Temperance Road
Franklin, Kentucky 42134

DOB: Franklin, Kentucky – October 15, 1952

EDUCATION
& TRAINING:

B.S. Agriculture Education
Western Kentucky University
1970-1974

M.S. Agriculture
Western Kentucky University
1976-1983

Rank I Educational Administration
Western Kentucky University
1992-1994

Ph.D., Educational Leadership
University of Louisville
1994-2010

PROFESSIONAL
SOCIETIES:

American Association of School Personnel Administrators
Society of Human Resource Managers
Kentucky Association of School Human Resource Managers
Association of School Administrators
Kentucky Association of School Administrators
Association of Supervision and Curriculum Development

NATIONAL
PRESENTATIONS:

BCSPA Annual Conference, Vancouver, BC Canada
MASA Annual Conference, Jackson, MS
AASPA Annual Conference, Hartford, CT